Causality and Cosmic Inflation

Tanmay Vachaspati* and Mark Trodden†

Department of Physics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-7079, USA.

In the context of inflationary models with a pre-inflationary stage, in which the Einstein equations are obeyed, the weak energy condition is satisfied, and spacetime topology is trivial, we argue that homogeneity on super-Hubble scales must be assumed as an initial condition. Models in which inflation arises from field dynamics in a Friedman-Robertson-Walker background fall into this class but models in which inflation originates at the Planck epoch, *eg.* chaotic inflation, may evade this conclusion. Our arguments rest on causality and general relativistic constraints on the structure of spacetime. We discuss modifications to existing scenarios that may avoid the need for initial large-scale homogeneity.

It is well recognized that an early inflationary epoch can explain several of the observed features of the present universe [1,2]. The remarkable homogeneity of the universe as measured by COBE, the flatness of the universe indicated by some of the current cosmic data, the distribution of structure, and the absence of magnetic monopoles may all be simultaneously explained by invoking about 60 e-folds of cosmic inflation. This remarkable fact has spurred considerable effort in building models that realize an inflationary phase of the universe. The goal of inflationary models is to explain how an assumed non-inflationary universe after the big bang develops into an inflationary universe at some epoch. Eventually, after some 60 e-folds, the universe must gracefully exit the inflationary stage and enter the radiation epoch of standard cosmology.

There exist alternative explanations for some of the cosmological observations that inflation so naturally explains. The distribution of structure may follow from topological defects [3]; the absence of magnetic monopoles from details of particle physics [4], or the interaction of domain walls and magnetic monopoles [5]. Other observed features of the universe are harder to explain by non-inflationary means. If the universe is indeed flat, it would be hard to explain this observation without invoking inflation. (It is known that certain inflationary models can lead to a non-flat universe, and so flatness is not a generic prediction of inflation but one of certain models.) Finally, the homogeneity of the universe is virtually impossible to explain without invoking inflation and this is a key compelling feature of the theory.

The ability of inflation to smooth out the universe on superhorizon scales is an effective mechanism to explain the observed homogeneity only if the inflation itself does not require violations of causality. This means that we must assume a pre-inflationary epoch of the universe from which a small patch of the universe underwent inflation entirely by causal processes. Note that causality dictates that the inflation must be "local". In other words,

any spacelike section of the boundary of the inflating region must not extend beyond the causal horizon of the pre-inflationary spacetime.

The question we address here is: under what conditions is it possible to have *local* inflation?

The embedding of an inflating region (not necessarily undergoing *exponential* inflation) within an exterior cosmology is constrained by the nature of matter in the universe. This is best seen by employing the Raychaudhuri equation for the divergence of a congruence of future directed, affinely parametrized null geodesics. This congruence is taken to be normal to a two dimensional sphere centered at the origin of coordinates and may be in- or out-going (i.e. directed towards or away from the origin of coordinates, respectively). Let us denote the tangent vector field to the congruence by N^a. Then the divergence θ is defined by

$$\theta = \nabla_a N^a \ . \tag{1}$$

The Raychaudhuri equation is:

$$\frac{d\theta}{d\tau} + \frac{1}{2}\theta^2 = -\sigma_{ab}\sigma^{ab} + \omega_{ab}\omega^{ab} - R_{ab}N^a N^b \tag{2}$$

where τ is the affine parameter, σ_{ab} is the shear tensor, ω_{ab} the twist tensor and R_{ab} the Ricci tensor. (We follow the conventions of Wald [6].) The shear tensor is purely spatial and hence its contribution to the right-hand side is positive. The twist tensor vanishes since the congruence of null rays is taken to be hypersurface orthogonal. Then,

$$\frac{d\theta}{d\tau} + \frac{1}{2}\theta^2 \leq -R_{ab}N^a N^b \tag{3}$$

If Einstein's equations hold then

$$R_{ab}N^a N^b = 8\pi T_{ab}N^a N^b \ , \tag{4}$$

and if the weak energy condition is satisfied ($T_{ab}\xi^a\xi^b \geq 0$ for any timelike vector ξ^a), then by continuity

$$T_{ab}N^a N^b \geq 0 \ . \tag{5}$$
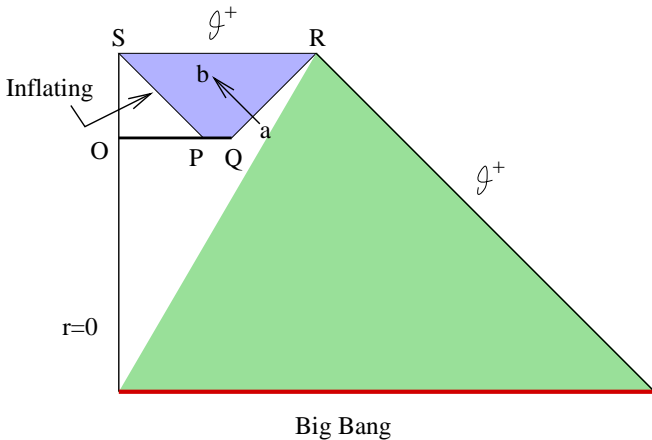
Putting these conditions together we obtain

$$\frac{d\theta}{d\tau} + \frac{1}{2}\theta^2 \leq 0 \ . \tag{6}$$

For our purposes, however, it proves sufficient to use the weaker condition

$$\frac{d\theta}{d\tau} \leq 0 \ . \tag{7}$$

Regions of a spherically symmetric spacetime in which the divergences of both in- and out-going rays, normal to spatial two dimensional spheres centered at the origin, are negative (positive) will be referred to as trapped (antitrapped) regions. Regions in which in-going rays have negative divergence (that is, are converging) but outgoing rays have positive divergence will be called "normal", since this is the behaviour in flat spacetime. Then the condition (7) says that a converging null geodesic cannot start to diverge prior to having reached the origin of coordinates, or *focussed*. In other words, in-going null rays cannot start out in normal regions and then enter an antitrapped region. This becomes the constraint in patching together an inflationary region in a background cosmology.

FIG. 1. A Penrose diagram for local inflation. The arrow denotes a future directed, radial, affinely parametrized null geodesic from the exterior spacetime into the inflating region. Shaded regions are antitrapped, unshaded regions are normal.

Consider a topologically trivial universe such as shown in Fig. 1. The universe starts out in a big bang and contains a normal region and an antitrapped region at distances larger than some distance that depends on the details of the cosmology. Now consider a patch of this region that starts to inflate. The patch is denoted by the horizontal line OQ, and has a physical size that we will denote by x_Q. The section OP of the line OQ denotes a spatial patch equal to the size of the inflationary horizon H_{inf}^{-1}. For inflation to occur, one assumes that vacuum energy must dominate over a region larger than the inflationary horizon distance, and so

$$x_Q \geq x_P = H_{inf}^{-1} \ . \tag{8}$$

Further, a straightforward calculation for ingoing null rays in spacetimes with the metric of the inflating region given by a flat FRW metric

$$ds^2 = -dt^2 + a^2(t)[dr^2 + r^2 d\Omega^2] \ , \tag{9}$$

yields

$$\theta = \frac{2}{a(t)}\left(H - \frac{1}{x}\right) \ , \tag{10}$$
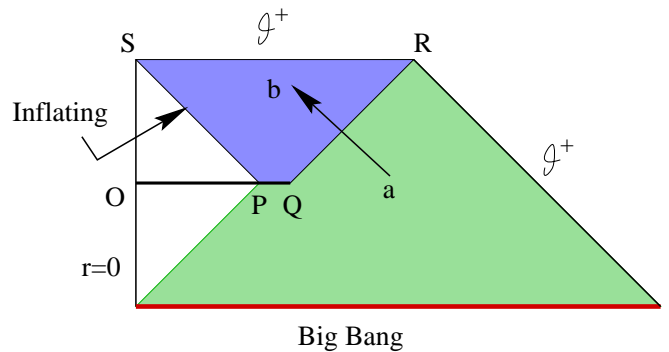
where, as usual, $H = \dot{a}/a$, and

$$x = a(t)r \tag{11}$$

where r is the coordinate of a null ray. Eq. (10) shows that the region with physical distance larger than H_{inf}^{-1} in the inflating region is antitrapped. Hence the region PQRS in Fig. 1 is antitrapped. The crucial question is: what are the allowed positions of the point P?

In Fig. 1 we show the situation where P is not located in the antitrapped region of the background cosmology. Then light rays such as shown in Fig. 1 (from point a to point b) can enter the inflating antitrapped region from the external normal region (unshaded in the figure). While the ingoing rays are in the normal region, θ is negative, but once they enter the antitrapped inflationary region θ must become positive. This is forbidden by the condition in eq. (7). Hence, we must conclude that the outer boundary PQR of the antitrapped inflating region must lie entirely inside the antitrapped region of the background cosmology as shown in Fig. 2. This is the key constraint on inflating spacetimes derived in this paper and, except for spherical symmetry, is independent of the background cosmology.

FIG. 2. A Penrose diagram for local inflation in which ingoing null geodesics that enter the inflating region emanate from antitrapped regions.

To appreciate the constraint, it is useful to think of the situation when the background cosmology is a flat Friedman-Robertson-Walker universe with a scale factor

$a(t)$. Then, the boundary of the antitrapped region of the background universe is given by

$$x_{FRW}(t) = H_{FRW}^{-1}(t) . \qquad (12)$$

Now, since the point P must lie within the background antitrapped region

$$x_P \geq x_{FRW}(t_P) , \qquad (13)$$

which yields

$$x_Q \geq x_P = H_{inf}^{-1} \geq H_{FRW}^{-1}(t_P) . \qquad (14)$$

This says that the size of the initial inflationary patch must be greater than the inflationary horizon, which must be larger than the background FRW inverse Hubble size at the time inflation starts. That is, the conditions appropriate for inflation to occur must be satisfied over a patch that is larger than the FRW inverse Hubble scale*

Note that the inverse Hubble distance, H^{-1}, can be different from the causal horizon (the distance light has propagated from the big bang). However, H^{-1} is in most cases still a large patch compared to length scales over which particle physics processes occur that can homogenize the universe. Also, for a flat, radiation dominated FRW cosmology, H^{-1} coincides with the causal horizon. We conclude that inflationary model building must assume homogeneity on super-Hubble scales. In this sense, inflationary models that attempt to obtain inflation within a background FRW universe cannot explain the homogeneity of the observed universe.

Our result is consistent with the result due to Farhi and Guth [10] who found that it is impossible to create an inflationary universe in the laboratory subject to the Einstein equations, the weak energy condition and the absence of singularities. On small enough scales in an expanding universe, it should be possible to ignore the background expansion and then the Farhi-Guth result should be applicable. This is consistent with our result since we find that the Hubble scale of the background spacetime provides a lower bound on the size of the inflating patch. If one admits the possibility of inflating false vacuum bubbles born at a singularity, the spacetime diagrams drawn by Blau, Guendelman and Guth

[11] show that the inflating region emerges from a white hole interior in which all two spheres are antitrapped. Then, once again, the boundary of the inflating region borders an antitrapped region.

In [12], Goldwirth and Piran numerically solved the Einstein equations together with a scalar field and found that inflation is obtained only if homogeneous initial conditions are assumed over a length scale that encompasses several horizons. (Similar numerical analyses were also performed by Kung and Brandenberger [13].) Our result generalizes and proves this numerical finding.

It is also worth pointing out that chaotic inflation [2] does not fall within the purview of our result since, in this model, inflation starts at the Planck epoch with homogeneity assumed on the Planck scale.

Hence we conclude that, within the conditions described above, local inflation is not possible. However, observations indicate homogeneity of the early universe on super-horizon scales and this needs an explanation. It is indeed possible that the initial homogeneity required by inflation - or even the homogeneity of the entire visible universe - occurred just by chance. Whether this is a satisfactory resolution of the observed homogeneity of the universe is largely a matter of personal taste and possibly anthropic considerations.

Let us now discuss the conditions under which local inflation can occur without assuming accidental homogeneity on large scales. The first possibility is that the weak energy condition (for diagonal, spherically symmetric fluid energy-momentum tensors, this amounts to assuming $\rho \geq 0$ and $\rho + P \geq 0$.) may be violated. For this one would need exotic forms of matter in the early universe. An attractive alternative is that quantum effects could give rise to effective violations of the weak energy condition. (However these violations are constrained by the Ford-Roman inequalities [14].) Whether quantum effects can be sufficient to lead to local inflation is an interesting question that has not yet been answered (an early related attempt was made in [15]). The second possibility is that the Einstein equations may be modified, leading to changes in eqs. (4) and (6). This is possible, for example, if we have a non-minimally coupled scalar field in the model. To us, this way out seems to be the best possibility especially in view of modern particle theories in which such scalar fields are abundant. A third possibility may be to have a topologically non-trivial background universe. Such universes have attracted significant attention recently [16] and should be investigated further. A fourth possibility is the one that occurs in topological inflation within magnetic monopoles as we discussed in detail in [17]. Here the inflation is manifestly local and causal but is preceded by a singularity or topology change (see Fig. 3). Conflict with our constraint is avoided because there are no null rays that enter the inflating region from the external region. The singularity or topology change plays the role of a mini big bang for the inflating spacetime,

---

*Based on homogeneous FRW cosmologies, a similar conclusion ($\dot{H}_{FRW} \leq 0$) was reached by Linde and collaborators [7–9]. Though leading to the same qualitative conclusion, there are differences in the final results since Linde and collaborators compare the Hubble scales for homogeneous universes at two different times while we are considering inhomogeneous universes and the Hubble scales are compared at the same time. However, as we also describe in this paper, the conclusion has been used as an argument against the viability of new inflationary models by these authors.

though is somewhat different in character from an FRW big bang since it is timelike. In any case, predictability in the inflating universe is lost because of signals that can emanate from the singularity or topology changing event. It is not possible to evolve to the inflating region from data on a spacelike hypersurface in the pre-inflationary epoch. Instead, initial data must be provided on a spacelike surface (Σ) within the inflating region.
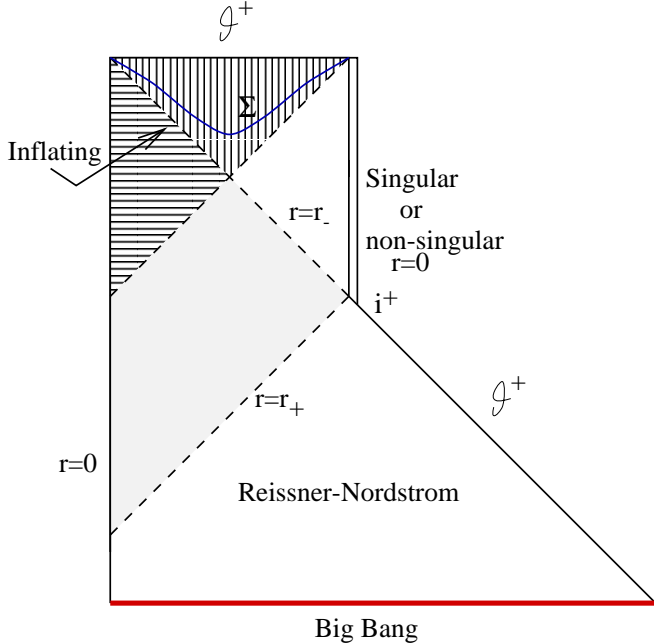

FIG. 3. A Penrose diagram for local inflation as in topological inflation with magnetic monopoles. Initial data must be provided on a spacelike hypersurface Σ entirely within the inflating region.

To summarize, we have argued that inflationary models based on the classical Einstein equations, the weak energy conditions, and trivial topology, require initial homogeneity on super-Hubble scales. Inflation with no requirements of initial large-scale homogeneity can be achieved with one or more of the following conditions: 1) violations of the classical Einstein equations, say due to non-minimally coupled scalar fields, 2) violations of the weak energy condition in the early universe, 3) non-trivial topology of the universe, 4) the birth of the universe directly into an inflating universe, that is, the absence of a pre-inflationary epoch, such as might occur in specific inflationary models, *eg.* chaotic inflation, and/or in the context of quantum cosmology [18–20]

[1] A. Guth, *Phys. Rev.* **D23**, 347 (1981).
[2] A. Linde, *Particle Physics and Cosmology* (Harwood, Chur, Switzerland, 1990).
[3] A. Vilenkin and P. Shellard, *Cosmic Strings and Other Topological Defects* (Cambridge University Press, 1994).
[4] P. Langacker and S.-Y. Pi, *Phys. Rev. Lett.*, **45**, 1 (1980).
[5] G. Dvali, H. Liu and T. Vachaspati, *Phys. Rev. Lett.* **80**, 2281 (1998).
[6] R. Wald, *General Relativity* (University of Chicago Press, 1984).
[7] See section 8.6 of [2].
[8] A. Linde and D. Linde, *Phys. Rev.* **D50**, 2456 (1994).
[9] A. Linde, D. Linde and A. Mezhlumian, *Phys. Rev.* **D49**, 1783 (1994).
[10] E. Farhi and A. Guth, *Phys. Lett.* **183B**, 149 (1987)
[11] S. Blau, E. Guendelman and A. Guth, *Phys. Rev.* **D35**, 1747 (1987).
[12] D. Goldwirth and T. Piran, *Phys. Rev. Lett.* **64**, 2852 (1990).
[13] J. Kung and R. Brandenberger, *Phys. Rev.* **D40**, 2532 (1989).
[14] L. H. Ford, *Phys. Rev.* **D48**, 776 (1993); L. H. Ford and T. Roman, *Phys. Rev.* **D51**, 4277 (1995).
[15] E. Farhi, A. Guth and J. Guven, *Nucl. Phys.* **B339**, 417 (1990).
[16] N. Cornish, D. Spergel and G.D. Starkman, *Phys. Rev. Lett.* **77**, 215, (1996).
[17] A. Borde, M. Trodden and T. Vachaspati, "*Creation and Structure of Baby Universes in Monopole Collisions*", gr-qc/9808069, to appear in *Phys. Rev.* **D** (1998).
[18] A. Vilenkin, *Phys. Lett.* **B117**, 25 (1982).
[19] J. Hartle and S.W. Hawking, *Phys. Rev.* **D28**, 2960 (1983).
[20] A. Linde, *Sov. Phys. JETP* **60**, 211 (1984); *Lett. Nuovo Cimento* **39**, 401 (1984).